Proprio

Latent Self-Scoring and Inference-Time Refinement for Physically Plausible Video Generation

A frozen video generator can score and improve its own videos

Mariam Hassan[1] Kaouther Messaoud[2] Wuyang Li[1] Alexandre Alahi[1]

[1] EPFL [2] Télécom Paris, IP Paris

Physical mistakes in generated videos

Check website for the video

Baseline

Proprio-refined

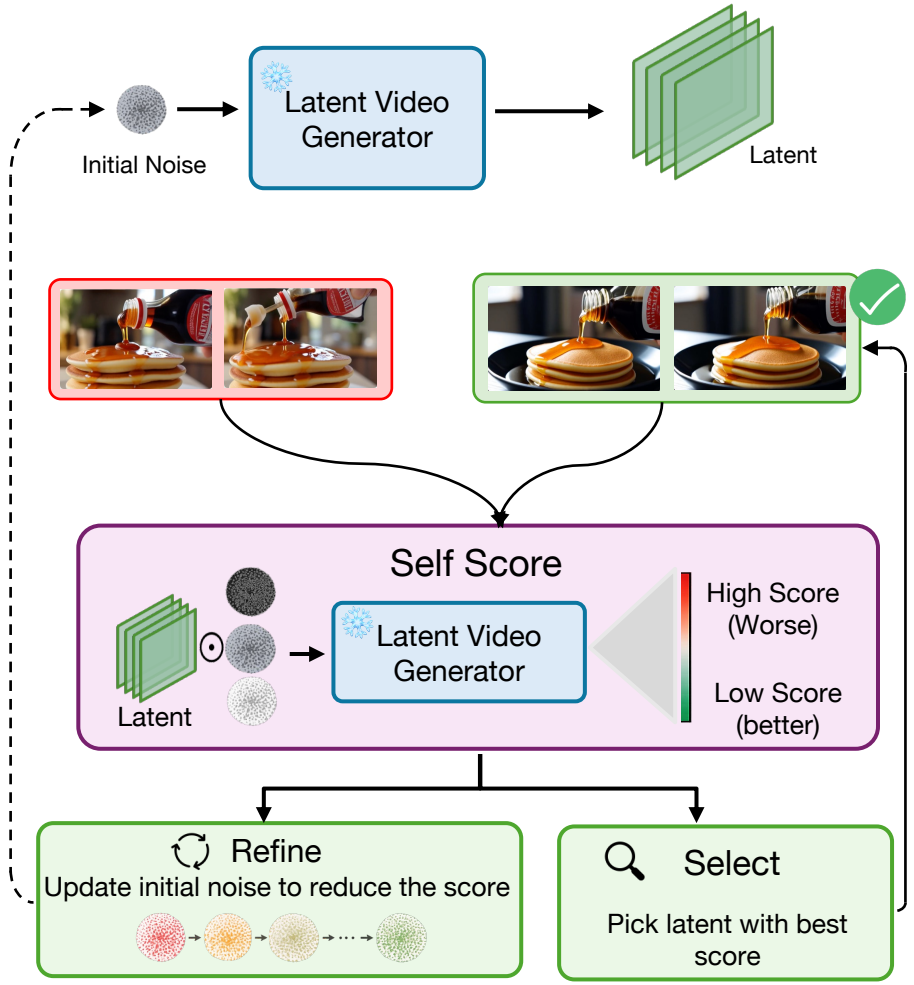


One model for generation, scoring, and improvement

1 Perturb

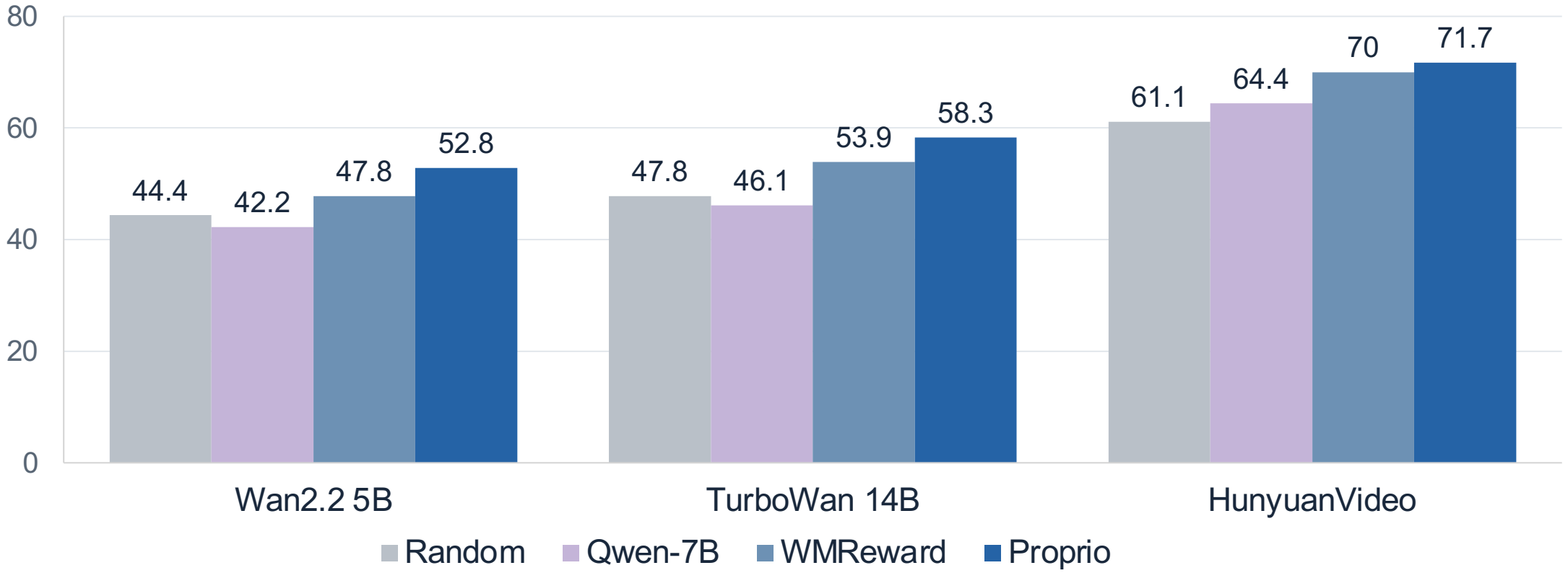
2 Measure

3 Score



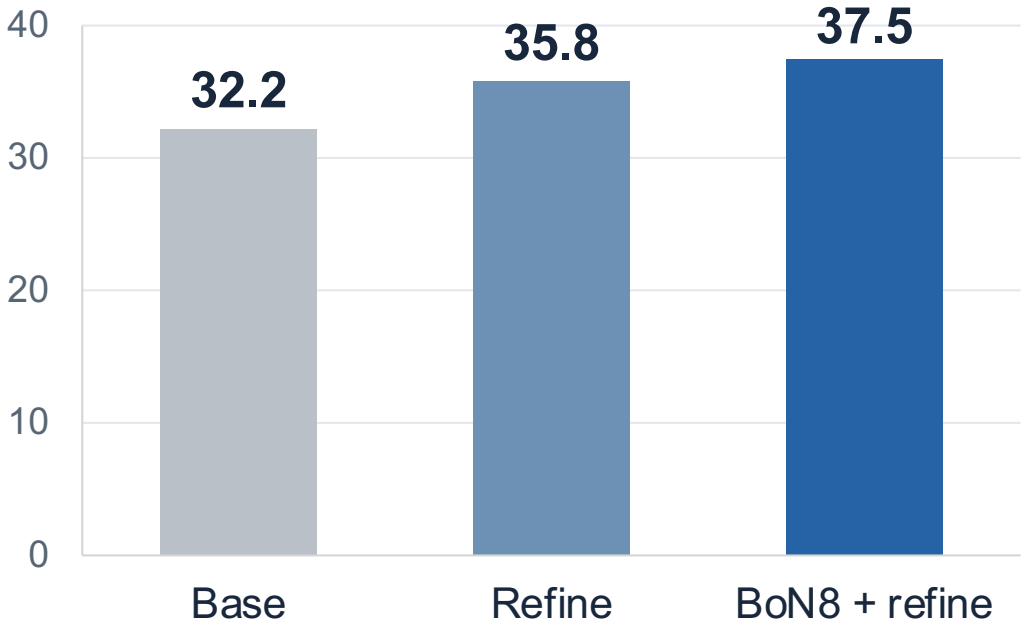
Self-scoring improves sample selection

VideoPhy2 hard split: physical commonsense accuracy (%)



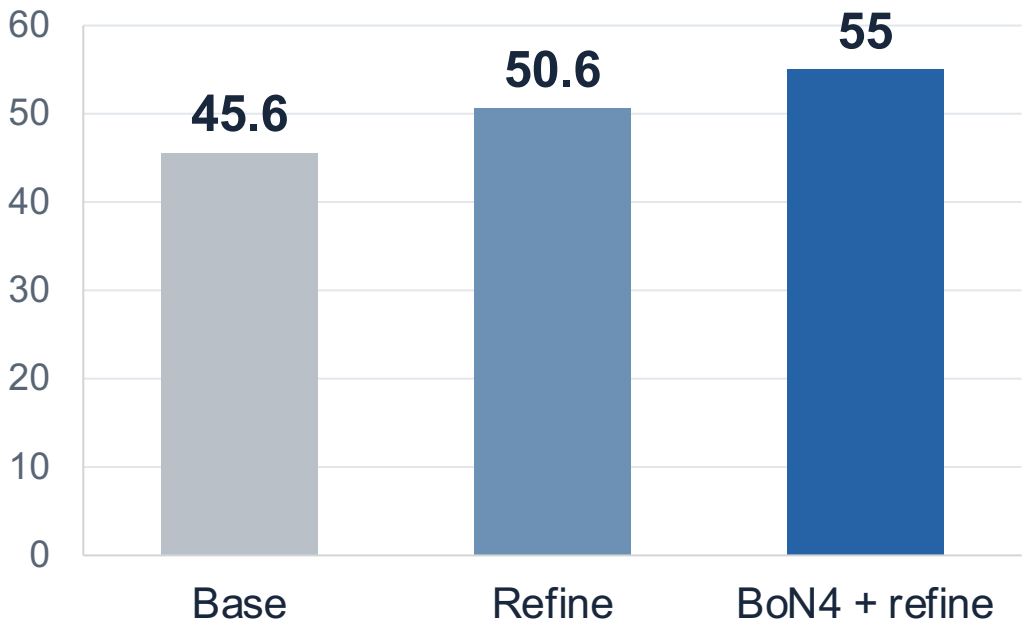
Search and refinement improve physical plausibility

Physics-IQ score



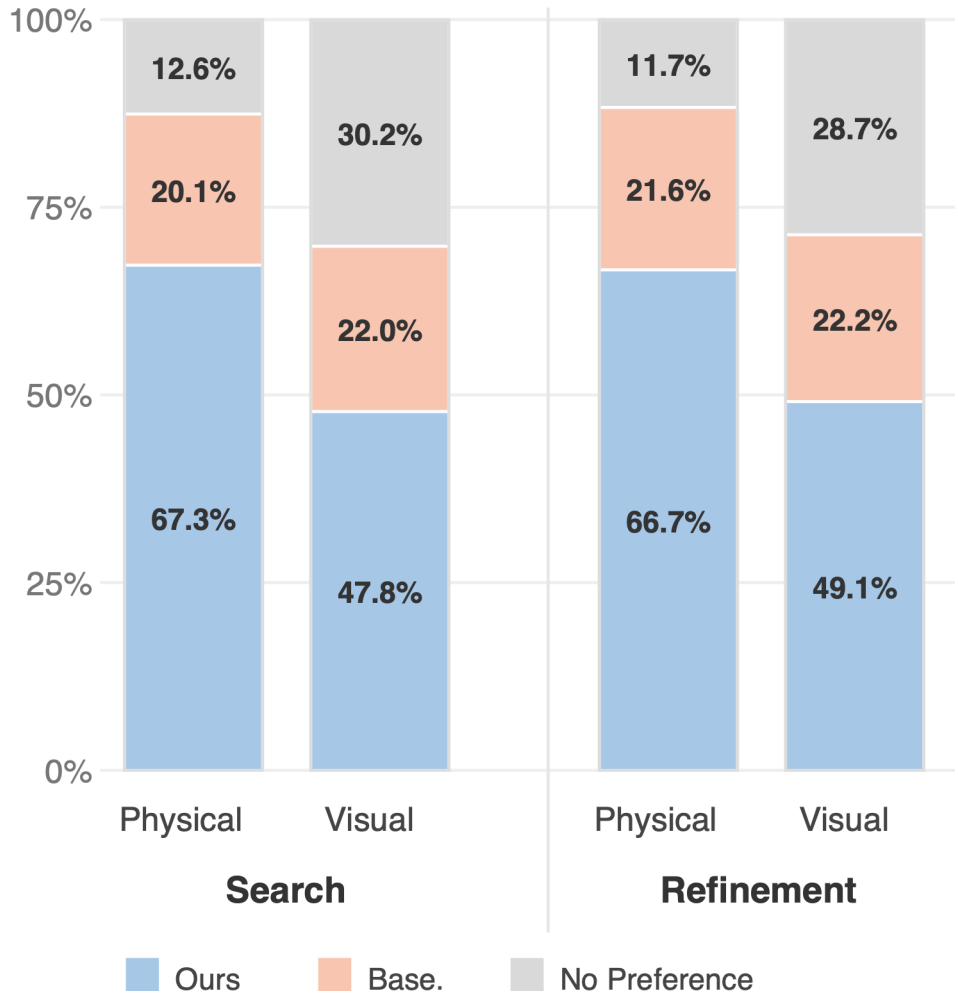
+16.5% relative to base

VideoPhy2 hard PC (%)



+20.6% relative to base

Human judgments support the self-scoring signal



A frozen generator can guide its own improvement

Baseline





Proprio-refined





One score for selection and refinement

More videos